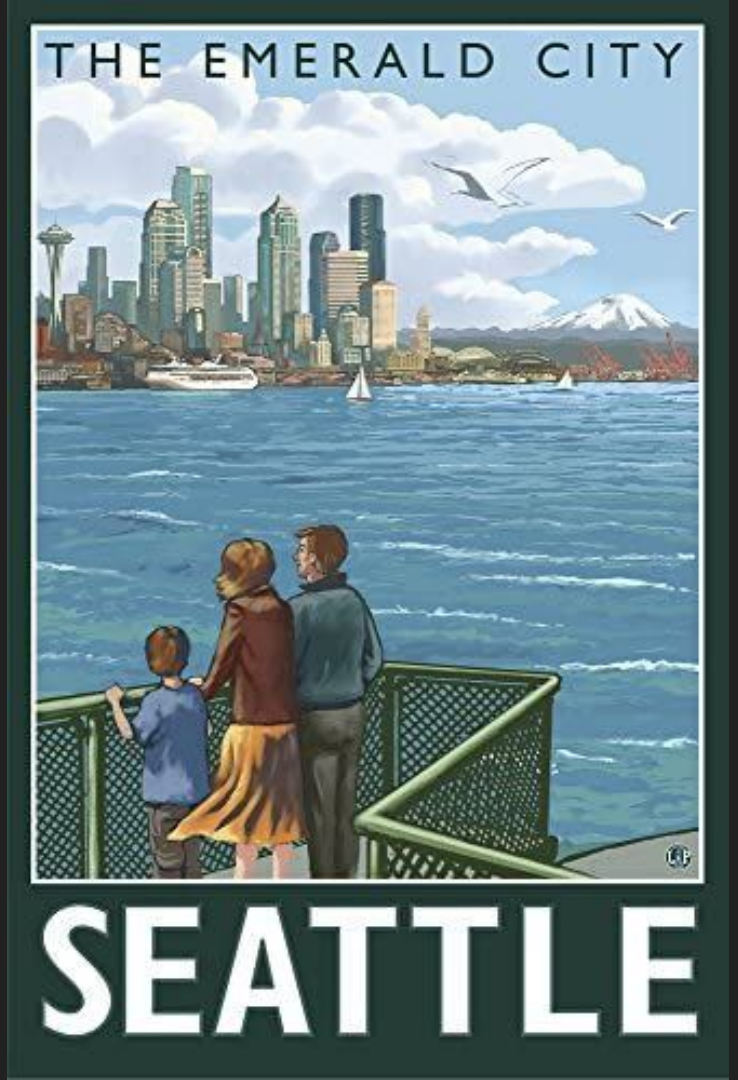
By : Ry AVERY, Mohammed ARAB, Mohamed Salim BENSAIDANI

CRIMES IN SEATTLE

Introduction

Seattle, Washington is widely known as The Emerald City hosting an estimated population of 755,000 over 142 sq miles of land. In 2020 Seattle was named the fastest growing city in America! Unfortunately, a large bustling city like Seattle is prone to a variety of crimes. From theft, fraud, property crimes and other various violent crimes, Seattle has seen its fair share of crime. This presentation will provide an overview of crimes that occur in Seattle, as well as an analysis of the crime statistics and trends. It will aim to inform those interested in what factors cause crime in American cities.



Project Overview

Goals

To understand the patterns, trends, and factors that contribute to crime in the city of Seattle.

- Understanding crime patterns would better inform public safety policies, initiatives, resources management, and support crime prevention efforts.

Ultimately, the goal of this project is to discover factors of crime rates and improve public safety in the city.

Summary

An initial look at crime statistics in Seattle show:

- Crime rates have been increasing in recent years.
- Certain areas of the city still experience higher crime rates than others.
- Main district has seen a significant increase in property crime in recent years.

Our analysis is about crimes rates in *King County* between 2017 to 2021.

Resources

- Seattle_PD_Arrest_Data_2017-2022_clean.csv
- census_data_final.csv
- Seattle_property_crimes.csv

Who will this inform

Crimes analysis in Seattle can inform a variety of stakeholders including government officials, law enforcement agencies, community organizations, and the general public. The information generated from crimes analysis can help identify crime trends, allocate resources more effectively, and develop targeted crime prevention and intervention strategies. Additionally, this information can be used to evaluate the effectiveness of existing policies and programs and make data-driven decisions to improve public safety.

Project questions?

What are the overall crime trends in seattle ?

What is the trend of crimes over the five years from 2017 to 2021?

How would a heat map of the crime rate for each zip code appear?

What are the overall personal crimes for Seattle for data period?

What is the correlation between demographic data and crimes data in Seattle?

Cleaning the Mess of Data Collected

- To analyze crimes in Seattle between 2017 and 2021 using Pandas, we obtained data on Seattle census data form those years, the number of property crimes that took place, and the public arrest reports that were filed during that time. We obtained this data from a publicly available crime database of the Seattle police department and through public data at census.gov/data/datasets.html.
- The CSVs we had to clean took a majority of our time spent on the project, as some CSVs contained upwards of 375,000 rows to sort through.
- In order to trim down the data we are working with we picked columns that would be useful in our analysis, converted columns into wanted data types (such as datetime 64 for dates), converted missing or unwanted values to NaNs and dropped them, and renamed columns that were ambiguous

Getting Zipcodes from Mess of Data Collected

- One of our larger datasets only contained latitude and longitude so we had to convert the coordinates into zip codes in order to use with the other datasets and find out more precisely where crime was taking place in Seattle
- Using the uszipcode library we imported the class SearchEngine and were able to reverse Geocode zip codes much faster than if we used an API key (search is a variable for SearchEngine())

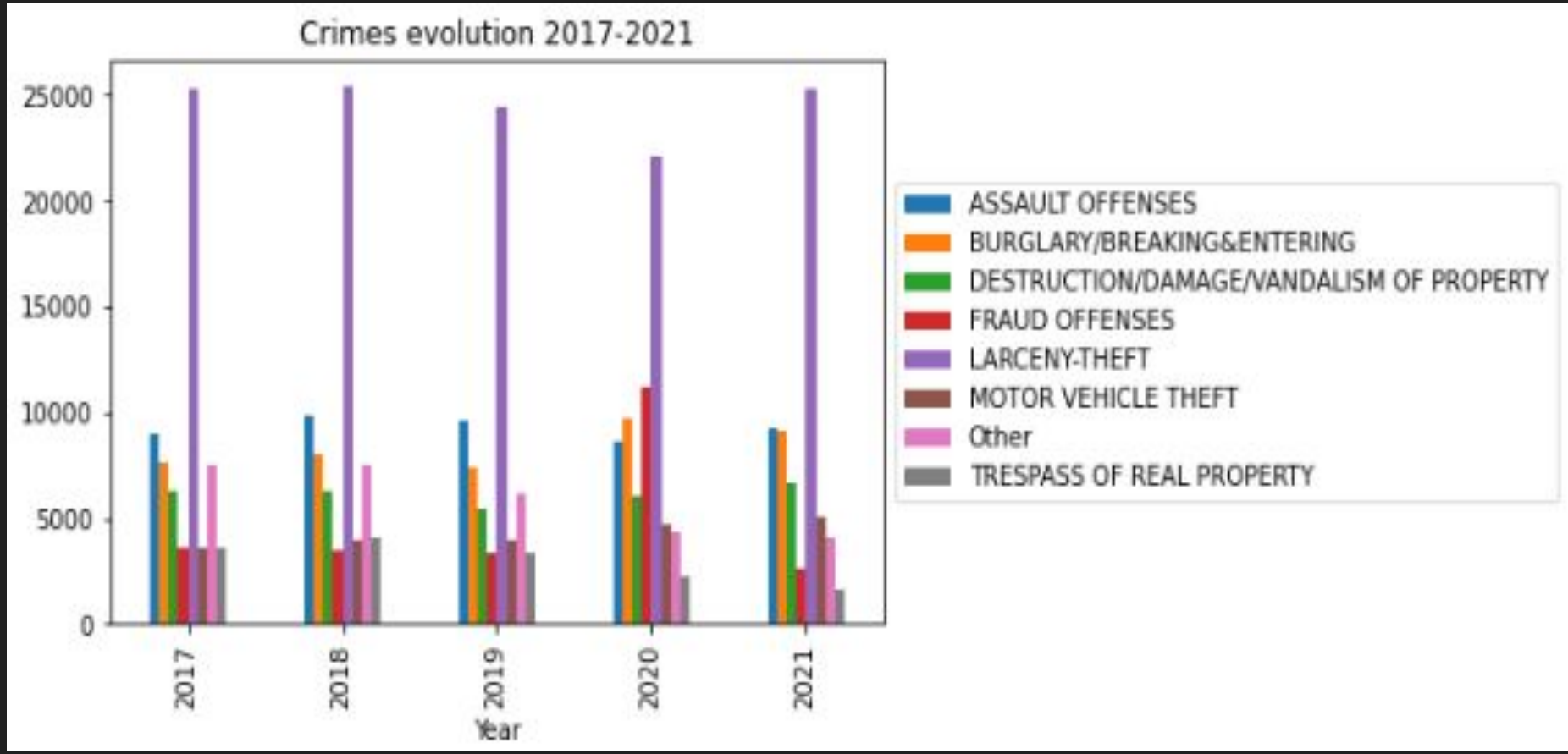
```
# Defined function to parse each rows ['Lat'] & ['Lng'] into zipcode search engine by coordinates,
# and extracted the zip from the results.
def get_zip(x):
    result = search.by_coordinates(lat=x['Lat'], lng=x['Lng'], returns=1)
    return result[0].zipcode

# For every row in spd_arrests_data_cleaned_df run the get_zip function, return the zip found to
# ['Zipcode'] series, print every 1000th record completed, and if an error come up skip that row.
for i, (idx, row) in enumerate(spd_arrests_data_cleaned_df.iterrows()):
    if (i % 1000) == 0:
        print(f'Working on {i}th record.')
    try:
        spd_arrests_data_cleaned_df.loc[idx, 'Zipcode']=get_zip(row)
    except:
        print(f'({idx},{row}) ran into error, Skipping...')
        pass
print('Zipcode Series Complete')
spd_arrests_data_cleaned_df['Zipcode']
```

```
0         98101
1         98125
2         98107
3         98101
4         98101
          ...
348541    98121
348542    98136
348543    98103
348544    98102
348545    98101
Name: Zipcode, Length: 348546, dtype: object
```

Preparing for Analysis on Data Collected

- Once we cleaned up the CSVs, we loaded them into DataFrames to perform operations and analysis on them. Some steps include:
  - Filtering the data: we used the df[df['column_name'] condition] syntax to filter the data based on a specific condition, such as only considering data from the desired time period.
  - Aggregating the data: we used the groupby method to aggregate the data by year.
  - Visualizing the data: we created various plots to visualize the number of property crimes in Seattle over time, such as a line plot, bar plot, and heatmap. We use libraries such as Matplotlib or hvplot.pandas to create these plots.
  - Performing statistical analysis: We performed various statistical tests and calculations to determine the trends and patterns in the data and we used a heatmap and correlation function

Crimes evolution 2017-2021

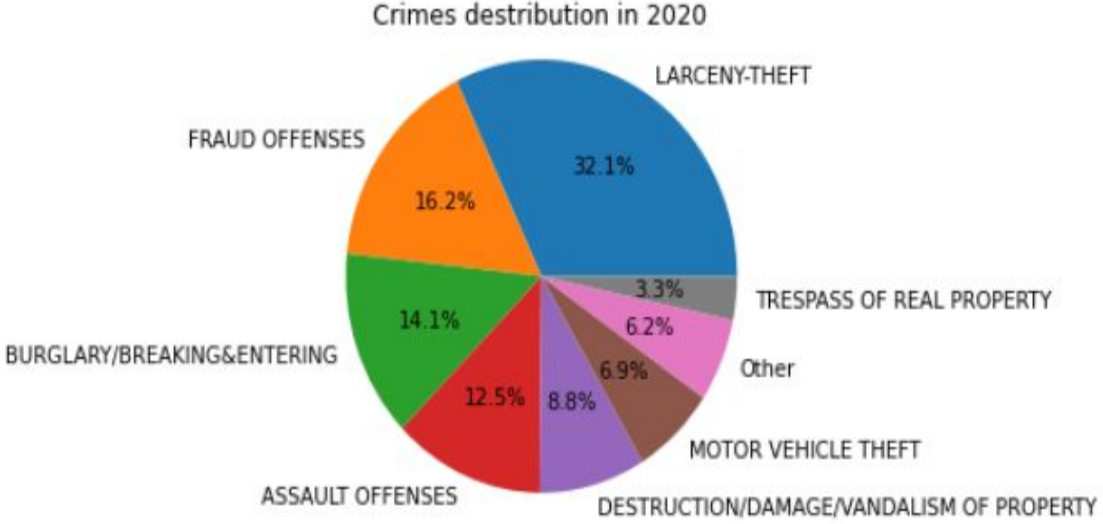


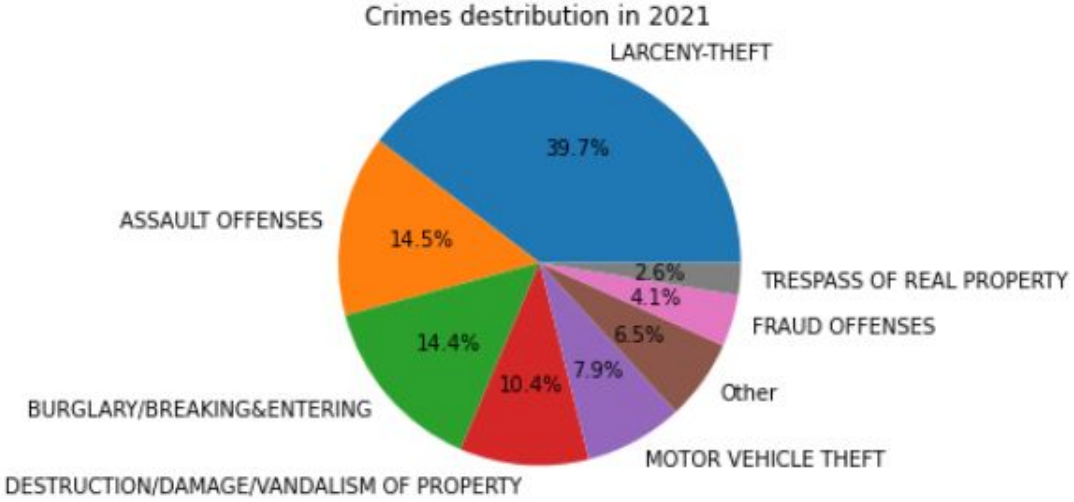
Violent Crime increased by 20%.
Property Crime increased by 9%.
Overall Crime in 2021 increased by 10% compared to 2020.

Offense Category	Offense	2021	2020	AMT CHANGE	% CHANGE
Violent	HOMICIDE	40	53	-13	-25%
	RAPE	253	270	-17	-6%
	ROBBERY	1,752	1,487	265	18%
	AGGRAVATED ASSAULT	3,295	2,656	639	24%
Violent Total		**5,340**	**4,466**	**874**	**20%**
Property	ARSON	224	171	53	31%
	BURGLARY	9,787	10,443	-656	-6%
	LARCENY-THEFT	26,733	23,188	3,545	15%
	MOTOR VEHICLE THEFT	5,305	4,912	393	8%
Property Total		**42,049**	**38,714**	**3,335**	**9%**
Grand Total		**47,389**	**43,180**	**4,209**	**10%**

*Reference : Police department of seattle

Distribution of crimes 2021 and 2020





Seattle Poverty Observations & Insights

Summary Statistics

Year	Mean Poverty Count	Median Poverty Count	Poverty Count Variance	Poverty Count Std. Dev.	Poverty Count Std. Err.
2017	3653.107143	3097.5	6.206818e+06	2491.348621	470.820634
2018	3507.071429	3104.0	6.384213e+06	2526.700101	477.501436
2019	3312.178571	2619.0	5.792639e+06	2406.790164	454.840588
2020	3171.892857	2506.0	4.570382e+06	2137.845287	404.014784
2021	3094.678571	2521.5	3.982410e+06	1995.597563	377.132491

From the table, there appears to be a decreasing trend in poverty count over the five years from 2017 to 2021. This is indicated by the following statistics

The poverty count steadily decreased over the five years, with the highest value in 2017 and the lowest in 2021

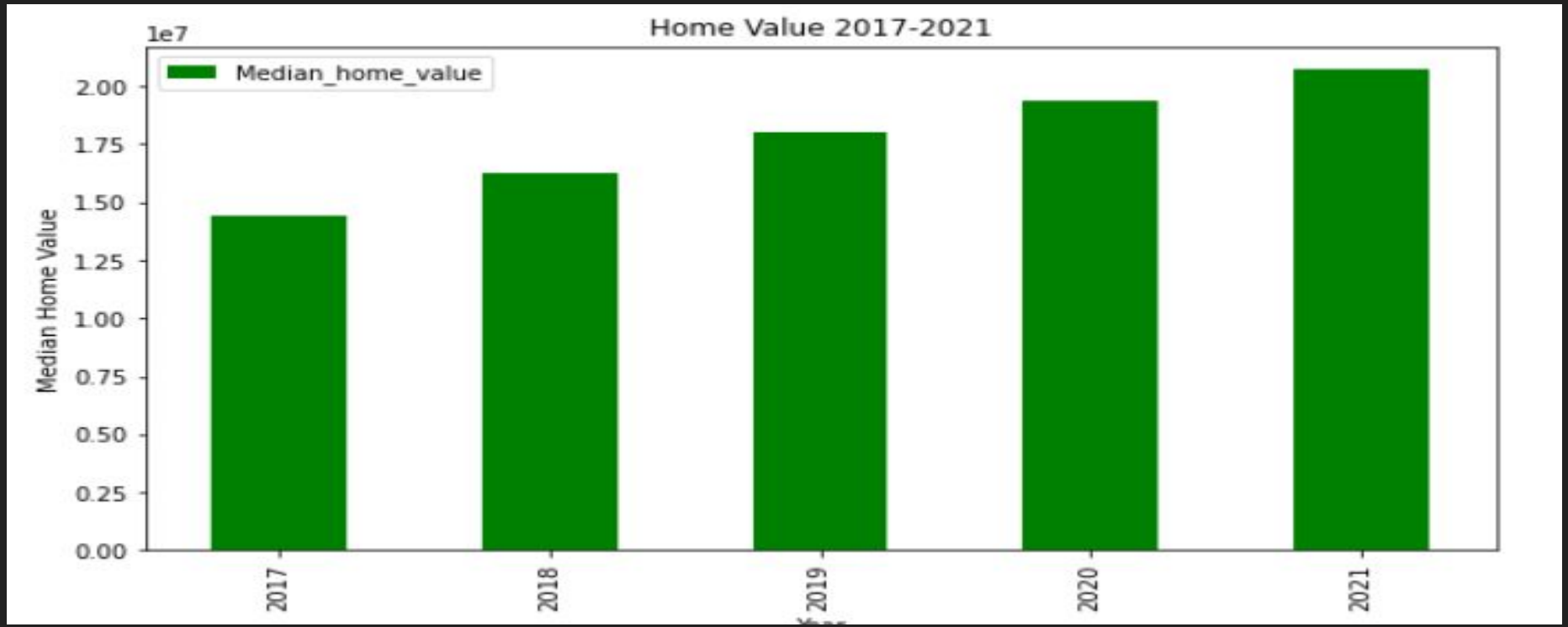
Seattle Property Crimes Observations & Insights

Summary Statistics

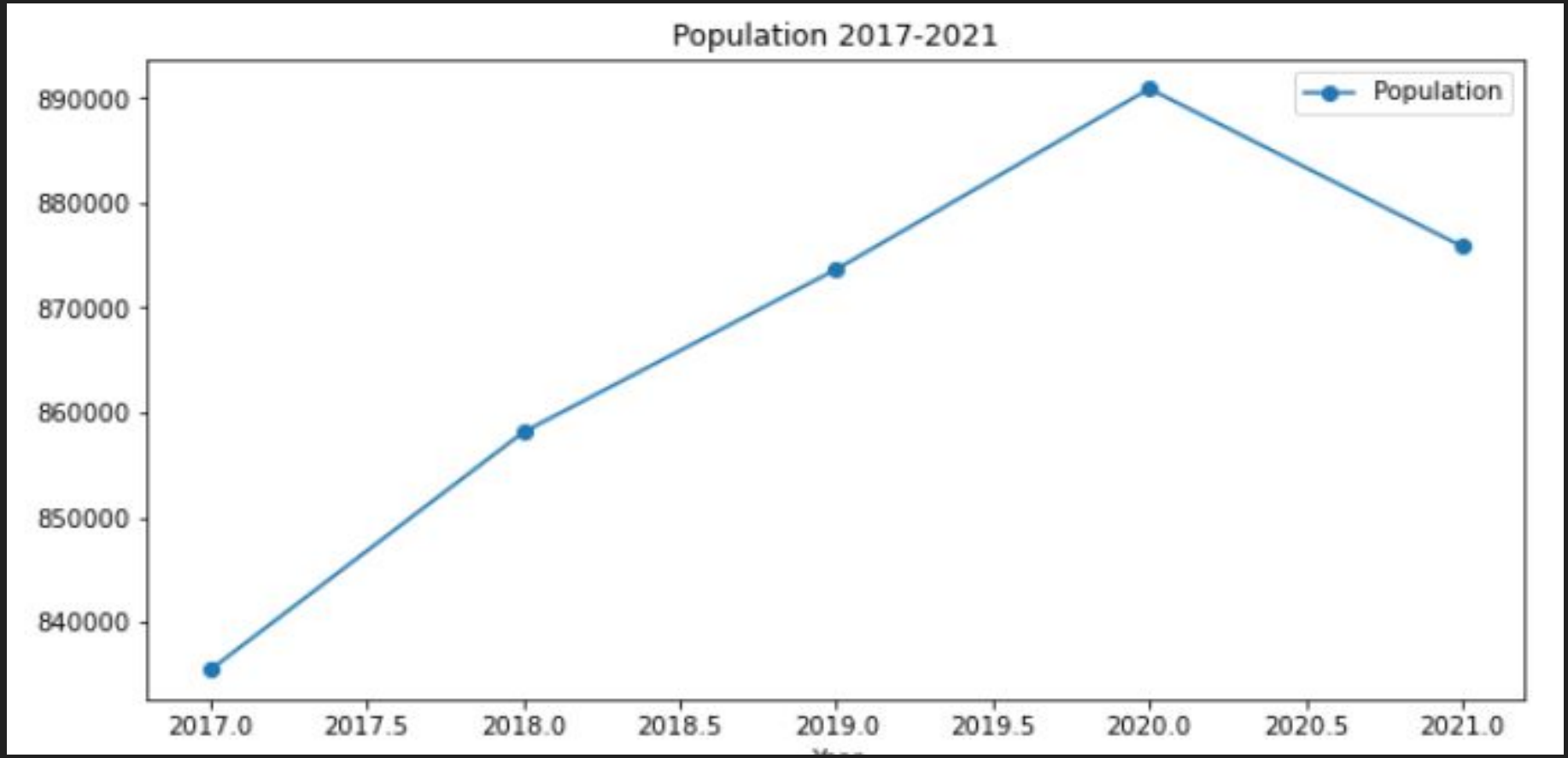
Year	Mean of Property Crimes	Median of Property crimes	Property crimes Variance	Property crimes Std. Dev.	Property crimes. Err.
2017	1351.071429	1228.5	710768.365079	843.070795	159.325404
2018	1325.642857	1245.5	683256.756614	826.593465	156.211482
2019	1193.428571	1147.0	564499.439153	751.331777	141.988360
2020	1348.678571	1274.0	738087.559524	859.120224	162.358461
2021	1385.678571	1318.0	782192.374339	884.416403	167.138990

From the table, there appears to be an increase in the mean and median of property crimes in 2020 and 2021, compared to the previous years

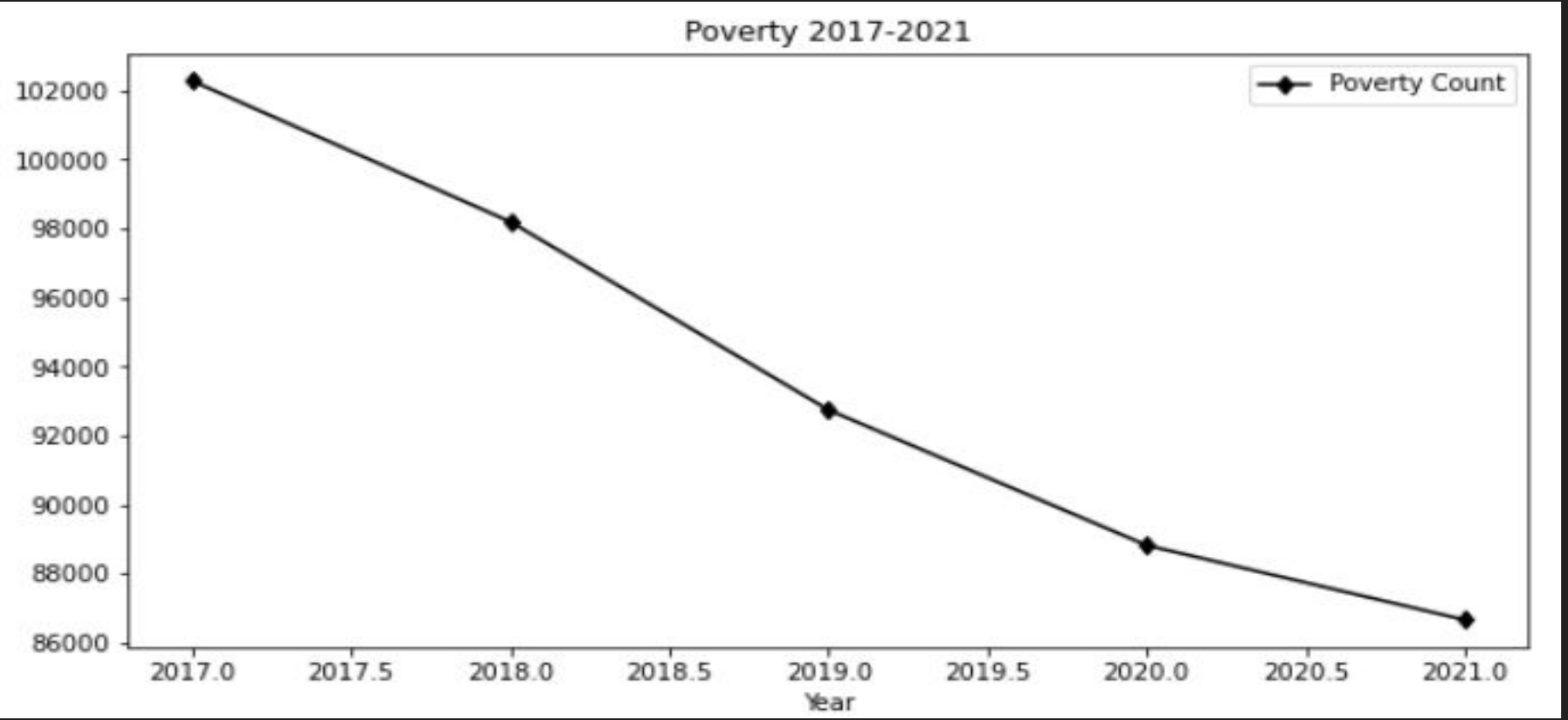
Evolution of Median Home Value over the years



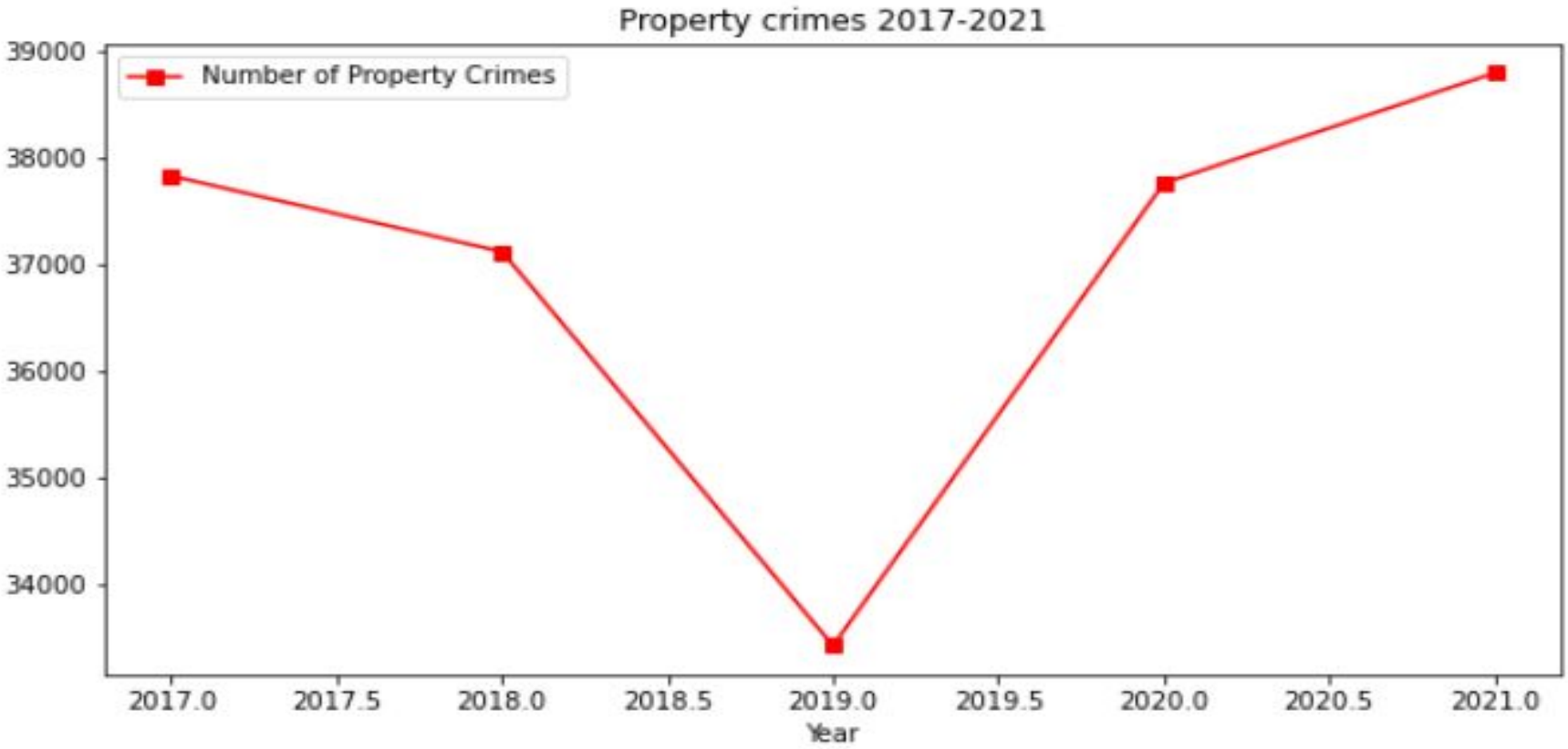
Increase of population



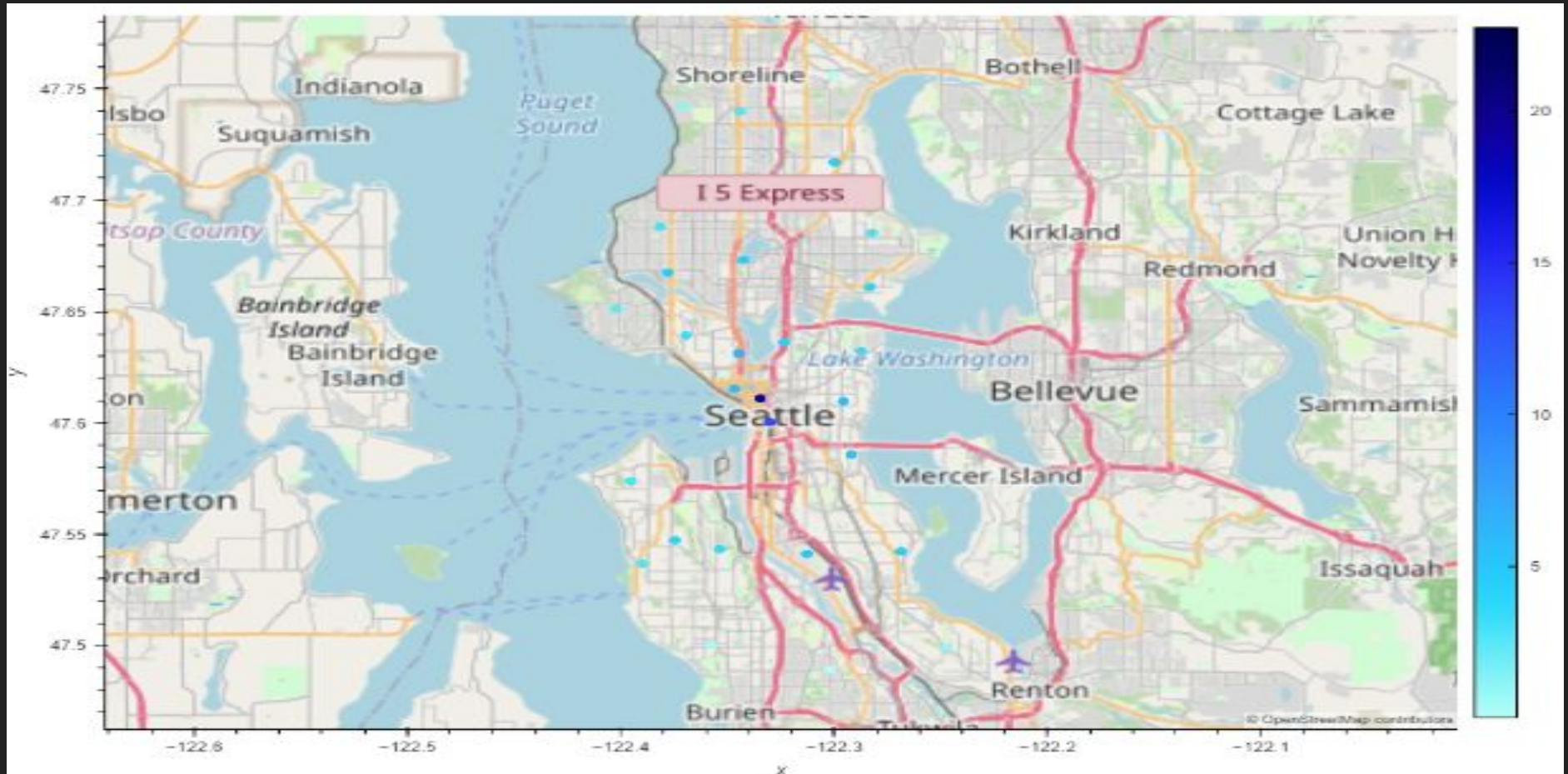
Decrease of poverty over the years



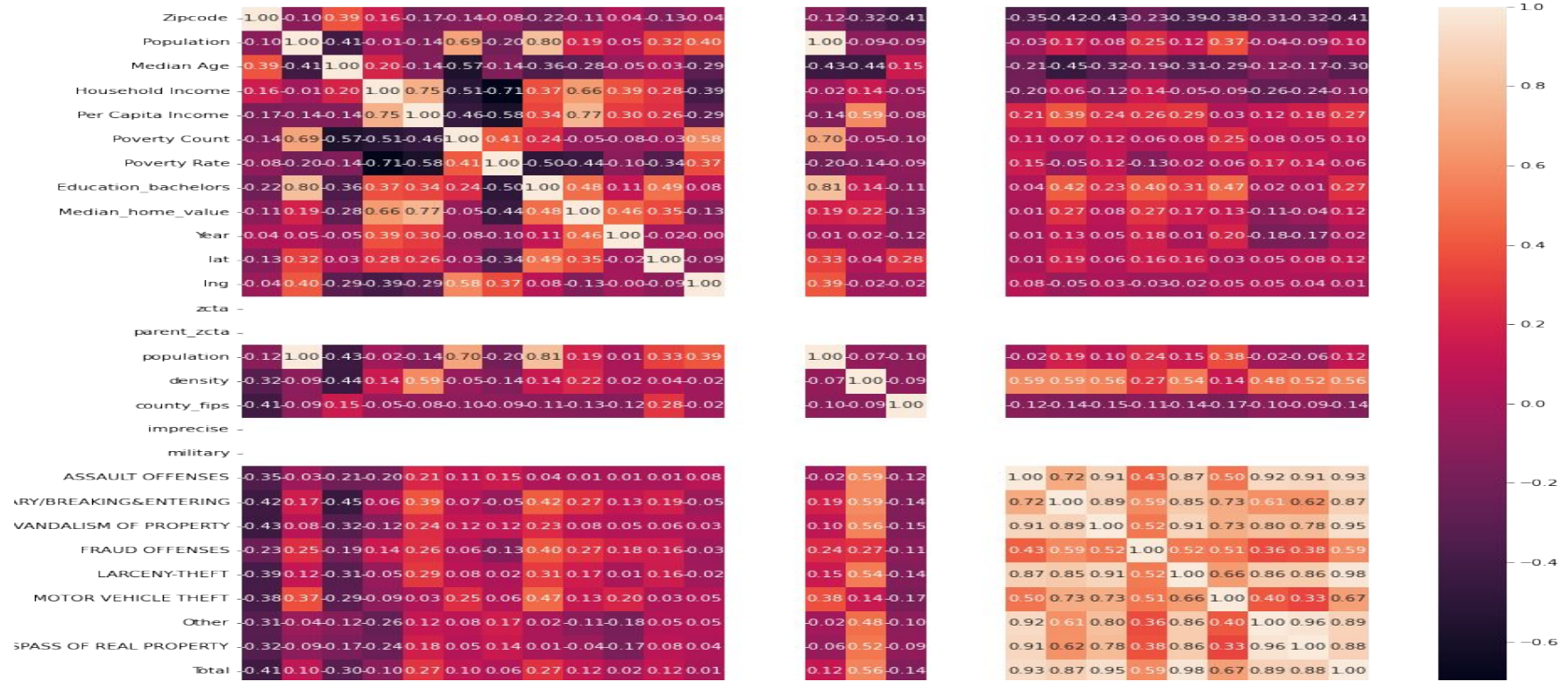
Evolution of property crimes over the years



Map of crimes rate in SEATTLE



The higher correlation displayed in this slide is the density by Zipcode and type of crimes. And we have the education bachelors count with fraud highly correlated.



Challenges we encountered

There is a several challenges encountered in crimes analysis in Seattle, some of them are:

1. Data Quality: Data collected from different sources may have errors or inaccuracies, which can affect the results of the analysis.
2. Data Availability: There might be missing or incomplete data, which can limit the scope of the analysis.
3. Data Integration: Combining data from multiple sources can be challenging and may require significant effort to ensure data consistency.
4. Data Privacy: Protecting the privacy of individuals involved in crimes is important, and thus, there might be restrictions on the data that can be used in the analysis.
5. Model Selection: There are many methods available for analyzing crime data, and selecting the right one can be difficult, especially when dealing with complex data.
6. Interpreting Results: Interpreting the results of a crime analysis can be challenging, especially when trying to draw meaningful conclusions from large amounts of data.

Conclusion

A crime analysis in Seattle can inform various stakeholders including the Seattle Police Department, city officials, community organizations, and residents. The analysis can provide valuable insights into the crime trends, patterns, and hotspots in the city. Based on these insights, the City officials and Seattle Police Department can allocate resources more effectively to reduce crime and improve community safety. Community organizations can use the analysis to identify areas where they can provide support and resources to improve public safety. Residents can use the analysis to make informed decisions about their own personal safety and the safety of their communities. Overall, a crime analysis can play an important role in improving community safety and reducing crime.